

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Papers in Behavior and Biological Sciences

Papers in the Biological Sciences

1989

Toward a Resolution of the Paradox of Aggressive Displays: II. Behavioral Efference and the Communication of Intentions

Alan B. Bond

University of Nebraska - Lincoln, abond1@unl.edu

Follow this and additional works at: <https://digitalcommons.unl.edu/bioscibehavior>



Part of the [Behavior and Ethology Commons](#)

Bond, Alan B., "Toward a Resolution of the Paradox of Aggressive Displays: II. Behavioral Efference and the Communication of Intentions" (1989). *Papers in Behavior and Biological Sciences*. 62.

<https://digitalcommons.unl.edu/bioscibehavior/62>

This Article is brought to you for free and open access by the Papers in the Biological Sciences at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Papers in Behavior and Biological Sciences by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

San Diego Natural History Museum, San Diego

Toward a Resolution of the Paradox of Aggressive Displays: II. Behavioral Efference and the Communication of Intentions

ALAN B. BOND

Received: December 14, 1987

Accepted: September 16, 1988 (G. Barlow)

Abstract

An inference from game-theory models of animal conflict is that adversaries should not inform one another about their level of aggressive motivation. This poses a paradox for the traditional ethological account of graded aggressive displays because it is usually assumed that the adaptive significance of these behavior patterns lies in their making such information available. To resolve the paradox, I propose that communication is only an incidental effect of displays, and that their primary adaptive function is regulation of the intensity of aggressive encounters through positive feedback on aggressive motivation, a process termed "behavioral efference." Evidence in support of this hypothesis is drawn from studies of human facial expression, aggressive catharsis, and operant conditioning of aggressive behavior. Implications of the hypothesis and suggestions for further work are discussed.

Introduction

Aggressive displays in vertebrates seldom consist of a single action pattern performed at a typical intensity (MORRIS 1957). Instead, in many species, the magnitude of several components of the display posture is continuously variable over a broad range (e. g. LORENZ 1966, on dogs and geese; LEYHAUSEN 1956, on cats; BROWN 1964, on Steller's jays). In others, aggressive communication involves a diverse collection of distinctive action patterns (ANDERSSON 1980). The latter case is typified by TINBERGEN's (1959) ethogram of twelve different aggressive behavior patterns in gulls or SERPELL's (1982) description of up to twenty aggressive action patterns in lorikeets.

One of the best-established tenets of ethology has been that this gradation in aggressive displays carries information about the motivational state of an animal

(LORENZ 1966; SMITH 1977). Where the behaviors are distinctive, they can usually be ordered in a sequence from low to high levels of aggressive motivation through their association with differing degrees of likelihood of attack or withdrawal in the contest (LORENZ 1981; TINBERGEN 1959; HINDE 1981). Similarly, continuously intergrading postures can generally be associated with a continuous variation in underlying motivational states (LORENZ 1966).

The second major component of the traditional ethological view is that aggressive displays serve as a substitute for combat, an alternative and less hazardous means of assessing aggressive superiority (TINBERGEN 1951). It seems essential to this view that resolution of conflicts by display and by fighting should yield similar outcomes: An individual that would have been defeated in actual combat should not generally be able to win through display alone. Thus, the central evolutionary function of aggressive display, in the traditional view, is to provide a truthful indication of the aggressive capabilities of the displaying individual. Given the empirical association between display behavior and aggressive motivation, it follows that the functional significance of complex, graded displays lies in their providing information to an opponent about the displaying animal's inner aggressive state.

This interpretation of aggressive display has recently been contested by investigators concerned with the evolution of display behavior. MAYNARD SMITH (1972, 1974) conceptualized animal conflict as a two-person game in which each participant chooses among a number of possible strategies, ranging from pure display to immediate attack, on the basis of their expected fitness payoffs. Each individual is seen as attempting to maximize its access to resources and to minimize its risk of active combat by manipulating the information available to its opponent.

The game formalism divides the attributes of the participants into two groupings: those that contribute to an underlying difference in relative fighting ability that would determine the outcome if the conflict were escalated to a physical attack, and those that reflect the individual's choice of display strategy. Through similar but independent lines of argument addressed to each of these groupings, game theorists have concluded that there is no selective pressure favoring truthfulness in aggressive displays with respect to either fighting ability or strategic intentions (MAYNARD SMITH 1974, 1982, 1984; DAWKINS & KREBS 1978; CARYL 1979, 1982 a, b). The latter argument, which is the principal concern of this paper, can be summarized as follows.

In a game between rational competitors, each participant's choice of strategy will be strongly influenced by the strategic decisions of his opponent: one is more likely to feel confident of attacking an opponent who is clearly unprepared to reciprocate. Revealing one's probability of attacking or willingness to continue the interaction ought to be avoided in these circumstances because it would allow the opponent to formulate effective countermeasures. Players in the aggression game should, therefore, avoid communicating their "intentions" to one another. By extension, motivational information should never be expressed in aggressive displays.

The use of the term "intentions" by game theorists is understandable, but unfortunate, because it appears to introduce irrelevant issues of consciousness and intentionality of communication. The moves of a chess master are certainly intentional, but it is hard to view the actions of an aggressive Siamese fighting fish in the same light. If we disregard the terminological inelegance, however, the game theoretic argument poses a serious problem for the traditional view of animal aggression. Simply stated, if the function of aggressive displays is to resolve a conflict over resources, and if the more aggressive individual will win a disproportionate share of such conflicts, why should it ever be adaptive to display a motivational level that is lower than one's opponent?

Aggressive display, by this argument, should evolve to a single action pattern that is consistently produced at maximum intensity, thereby providing no information about differences in motivation (MAYNARD SMITH 1984). The occurrence of complex, graded aggressive displays therefore constitutes a paradox (DAWKINS & KREBS 1978). If they are truthful reflections of the animal's internal state, why has revealing such information not been selected against? If they are not truthful indicators of motivation, why should they exhibit such remarkable diversity (ANDERSSON 1980; SERPELL 1982)?

The response of many ethologists has been to argue that graded signals do occur and are sufficiently predictive of the future behavior of the displaying animal, provided that one correctly interprets the conditional and probabilistic nature of the message (HINDE 1981; SMITH 1986; BARLOW & ROWELL 1984; VAN RHIJN 1980). Hence, displays do, in fact, provide valid information about the displaying individual. The implication is that the paradox is some sort of artifact of the game theory approach or the result of overlooking additional factors, such as the occurrence of repeated encounters between individuals that can remember and recognize one another (VAN RHIJN 1980; VAN RHIJN & VODEGEL 1980).

This does not, however, fully address the source of the problem. The paradox owes nothing to game theory, as such. The theorists simply encountered it in the course of developing a rigorous model for the evolution of displays. The importance of the game-theoretic argument is in revealing that the two central features of the traditional ethological account of aggression — that graded displays present valid information about the internal states of the displaying animal and that the display of such information is, in fact, the central function of the behavior — are fundamentally incompatible.

In a companion article (BOND 1989), I have suggested modifying the first assumption to allow for deception in the communication of fighting ability. This argument of "optimal deceit" cannot constitute a complete answer to the paradox, however. It addresses only components of fighting ability; its application to motivation or "intentions" is, at best, tangential (MAYNARD SMITH 1982; CARYL 1982 a).

To resolve the paradox, we must address the second assumption, that a central function of aggressive displays is the communication of motivational states (e.g. ANDREW 1963). One logical alternative is that displays *can* be used as predictors of subsequent actions and do, therefore, provide truthful reflections of

aggressive motivation or intent, but that truthful communication is only a side-effect. The principal adaptive function of graded aggressive displays, I wish to argue, is regulation of aggressive arousal.

The Argument for Behavioral Efference

The adaptive significance of aggressive behavior is, presumably, that more aggressive individuals often enjoy a greater access to limiting resources (ARCHER 1988). Aggression can also be dangerous, however: An animal that engages in active combat runs a significant risk of injury, even if it wins the encounter (GEIST 1974). In species that make use of aggressive displays, escalation to active combat is often unnecessary because many conflicts can be resolved by display alone (TINBERGEN 1951; DAWKINS & KREBS 1978). Individuals that resort too readily to physical attack, therefore, will fight many battles that need not have been fought and run a significantly higher risk of injury.

Rapid escalation can also be damaging to inclusive fitness. Stimuli that release aggression are presented by mates and offspring, as well as conspecific opponents (LORENZ 1966). Innumerable mechanisms provide protection against mate aggression, including individual recognition, submissive behavior, and bonding rituals (LORENZ 1966). All of these mechanisms require time, however, time for the mate to respond, time for the recognition to take effect. Individuals that attack apparent opponents immediately, with few preliminaries, may have difficulty in maintaining stable pair-bonds.

Although there may be a general benefit to being aggressive, therefore, the risk of injury to oneself and one's mate and offspring appears to be high enough that rapid escalation ought to be selected against. This is reflected in common parlance in the disapproval of individuals who cannot "hold their tempers," or who tend to "fly off the handle." It seems reasonable, therefore, to expect selection for a resting level of aggressive arousal that is considerably below the maximum, as well as for behavioral mechanisms that regulate the rate of arousal during an interaction.

In animals with complex, graded displays, aggressive interactions are characteristically escalated in that low-intensity behavior patterns are progressively replaced by higher-intensity ones (ARCHER 1988 cites several examples). Given the evidence, cited earlier, that such behavior is indicative of different levels of aggressive motivation, an escalating interaction indicates a graded arousal process, a gradual increase in aggressive motivation over the course of the encounter.

The adaptive significance of this behavior is that by postponing physical attack a graded arousal process increases the likelihood that contextual factors, such as submissive signals or individual recognition, can intervene and halt the interaction, thereby reducing the frequency of unnecessary combat and preventing injury to mates or offspring. This argument is similar to SCHERER's (1985) suggestion that requiring an individual to build up a particular level of arousal prior to acting, "allows a reevaluation of the eliciting antecedent event . . . and an evaluation of the likely consequences of the behavioral response alternatives."

As a corollary to this interpretation, the selective value of postponing attack should be greater in species with more formidable weapons. We might, therefore,

expect a relationship between the size of the repertoire of aggressive displays in a given species and the potential for causing injury. Such a relationship is beautifully illustrated in SERPELL's (1982) work on lorikeets, which are small, aggressive parrots in the genus *Trichoglossus*. In a survey of the display repertoires of nine species, Serpell found that the larger, heavier-billed species, which are capable of inflicting serious injuries, had repertoires of up to twenty highly ritualized display behaviors. In contrast, smaller species with a less powerful bite had repertoires of as few as five behaviors and a much higher tendency to escalate to physical attack (SERPELL 1982).

We must then ask why escalation is characteristically progressive. What causes the consistent increment in aggressive arousal? Clearly, a principal source of stimulation is provided by the opponent. DAWKINS & KREBS (1978) have likened aggressive interactions in species with graded displays to an auction, in which two individuals bid against each other at progressively higher levels, up to the point at which one of them withdraws (a similar analogy is proposed by DOW et al. 1976). It seems unlikely, however, that the behavior of the opponent is the sole source of aggressive stimulation. Many species will show intense, escalated attacks on mirror images, in spite of the fact that the "opponent" in these cases is incapable of making a higher bid (GALLUP 1968). Escalation commonly occurs even in interactions with inanimate models (HOGAN & ROPER 1978).

One means of accounting for these effects is the hypothesis of "behavioral efferece," of positive feedback from the expression or display of aggressive tendencies to the internal motivational state. Display is, in this view, not merely indicative of a particular level of motivation, but actually instrumental in achieving and sustaining it. The characteristic escalation of aggressive interactions is, then, simply a reflection of the positive feedback: Higher levels of aggression are commonly attained as a consequence of the animal's having previously performed lower-level aggressive displays.

The concept of behavioral efferece thus provides a means for directly regulating the rate of intensification of an aggressive interaction. Threshold levels of aggressive arousal are linked to specific display behavior, with each display in a graded series requiring a successively higher level of arousal in order to perform it. Performance of an aggressive display then increases the level of arousal, thus enabling the animal subsequently to produce an even more intense display. Unless other factors intervene, such as capitulation or withdrawal of one of the participants, the motivational and behavioral increments will cascade up the intensity scale until a level of arousal is reached that is compatible with physical attack.

Since aggressive interactions do not invariably accelerate to active combat, termination of the process must commonly be brought about by additional, inhibitory factors. There are indications of inhibitory feedbacks of aggressive performance on motivation: aggression appears to "sate" (POTEGAL & TEN BRINK 1984; BAENNINGER 1966). This appears to be true of aggressive display, as well as physical combat (HOGAN & ROPER 1978). The principal source of inhibitory regulation, however, is probably the behavior of the opponent.

It is in this context that the traditional account of displays as mixtures of aggression and fear, of arousal and inhibition, comes into play (HINDE 1981). The effect of the opponent's displays would depend on the level of aggressive motivation to which they corresponded. Displays indicative of higher arousal might be inhibitory, as would signs of submission or subordinate status. Displays indicating an equivalent or somewhat lower level of motivation, on the other hand, could serve to accelerate the rate of arousal (ARCHER 1988).

Displays do, thus, serve to communicate an opponent's intentions, but the communication is, in OTTE's (1974) terms, an "effect," rather than a "function" of the behavior. In the behavioral effference model, graded aggressive signals are the causal equivalent of "intention movements." Intention movements are not, in general, selected for communication, but are rather the mechanical preliminaries to some subsequent act (MOYNIHAN 1982). The characteristic crouch and partial wingspread in alarmed birds communicates their intention to fly, but the behavior also provides the necessary physical preparation for the leap that will make them airborne. Similarly, aggressive displays may incidentally communicate an animal's probability of attack, but only as a result of their function in increasing aggressive arousal.

If behavioral effference exists, why should it be evolutionarily stable? MAYNARD SMITH's (1982; 1984) argument concerns a hypothetical mutant that always displays the most extreme level of aggression, no matter what its actual intentions are, and thereby wins a disproportionate share of its aggressive encounters. What is to prevent such a mutation from spreading throughout the population? The reason may be that a mutant could not simply acquire the ability to produce displays without the corresponding level of motivation: it would also have to develop an alternative means of regulating aggressive interactions. Otherwise, decoupling behavior from arousal would free aggressive actions from the control provided by the graded arousal process and increase the risk of unnecessary injury.

An illustration of the difficulties involved in selecting for concealment of intentions is provided by the expression of the emotions in humans. The evolution of human language has provided a secondary channel for the communication of intentions and motivation, a channel that is not directly linked to the regulation of arousal. In consequence, it is relatively easy for humans to provide false verbal information, to declare "I am very angry" or "I love you" in the absence of any corresponding motivational state. It is, however, extremely difficult for most people to lie *successfully* about motivations, because the older, nonverbal components of behavior that are still tied into the original control mechanism provide contrasting signals that give away one's true intent (EKMAN 1985). Even professional actors, when asked to simulate a specified emotional state, can do so effectively only by deliberately recalling an incident in their past that evokes an appropriate emotional accompaniment (EKMAN 1985).

The Evidence for Behavioral Effference

Behavioral effference is scarcely a novel idea. Idioms such as "working oneself into a rage" (cf. LORENZ 1966, p. 154) convey a common acceptance of the

need for self-arousal prior to a physical attack. Familiar examples include the "warm-up" activities performed by athletes or warriors before a contest, in which they engage in aggressive displays, vocalize, and build their hostility in the absence of opponents, so as to be fully prepared when the moment of combat arrives (POTEGAL 1979).

Direct experimental evidence of the influence of performance on motivation is remarkably rare in the ethological literature, however. There are some data available from studies of courtship behavior, in which the stimulatory or potentiating effects of display performance have long been recognized (e. g. MORRIS 1956). In a particularly elegant experiment, WILZ (1970) was able to show that performance of the "creeping through" display by male sticklebacks was instrumental in switching their own predisposition from aggression to courtship.

The absence of similar results on the effect of aggressive displays may simply indicate that they have not been looked for. The traditional paradigm has been that motivation accumulates continuously until the behavior is "released" by an appropriate stimulus (LORENZ 1950). This viewpoint appears to have focused primary attention on the role of the stimulus, rather than the response, in aggression. For example, HEILIGENBERG (1965) and LEONG (1969), who conducted some of the best-known and best-regarded experiments on aggressive motivation (HUNTINGFORD 1980), concerned themselves exclusively with the effect of the stimulus models they presented and apparently did not consider the effects of the display behavior that the models elicited. Indirect evidence of behavioral efference in aggression can, however, be adduced from a variety of different sources, including studies of human facial expression, of aggressive catharsis, and of the operant conditioning of aggressive behavior.

Studies of Facial Expression

The theory of "facial efference" asserts that the experience of emotion derives from feedback from facial expressions. The roots of the idea can be traced to DARWIN's 'Expression of the Emotions in Man and Animals.' In his summary chapter, he stated the hypothesis in unequivocal terms:

"The free expression by outward signs of an emotion intensifies it . . . He who gives way to violent gestures will increase his rage; he who does not control the signs of fear will experience fear in a greater degree . . . These results follow partly from the intimate relation which exists between almost all the emotions and their outward manifestations . . . Even the simulation of an emotion tends to arouse it in our minds" (DARWIN 1872, p. 366).

In recent years, this perspective has provided the basis for a substantial body of theory and experiment on the functional significance of emotional expression (TOMKINS 1962; IZARD 1977; EKMAN et al. 1983; ZAJONC 1985). Although the exponents of this theory are sharply divided on the details of the intervening mechanism, they all concur that the clearest causal relationship for the intensity of affect, or emotional experience, results from feedback from its expression. This is especially true of the species-typical, innate expressions that, in primates, are generated primarily by the face (EKMAN 1980).

"Emotion" is a multifarious phenomenon, with cognitive, as well as physiological and affective aspects (COLLIER 1985; HINDE 1985). The continuity between emotional and motivational systems has commonly been emphasized, however (TOATES 1986). This is particularly true for nonhuman species, where the cognitive aspects of emotion are less significant (e. g. ANDREW 1963; WEINRICH 1980; TOMKINS 1984; PLOOG 1986). The equation of animal displays with human emotional expression was the central theme of DARWIN's (1872) work and has been broadly accepted in contemporary ethology (e.g. EIBL-EIBESFELDT 1972). The relationship between affect and expression in human emotions is, thus, directly applicable to the issue of the functional significance of display behaviors.

The weight of the experimental evidence strongly favors some form of the facial efferece hypothesis (LAIRD 1984; COLLIER 1985). In one protocol, subjects are instructed to contract and relax facial muscles that, in combination, produce an expression associated with a particular emotional state. They are given a plausible excuse for the experiment, usually electromyographic recording, to misdirect their attention, and no reference is made to the emotional content of the resulting expression. Subsequent evaluation of measures of both self-report of emotional experience and physiological arousal has indicated that induction of the facial expression causes the experience of a corresponding emotional state. This effect has been replicated in a large number of studies, even when data from subjects who later indicated an awareness of the emotional content of the induced expression were excluded (LAIRD 1984).

A second protocol involves exposing subjects to emotionally evocative stimuli, generally film images. An observer is present who can watch the subjects' expressions, but cannot see the screen. The subjects are instructed to exaggerate or suppress their expressive reactions, supposedly in order to deceive the observer. Such studies have also shown significant effects on both self-report measures and physiological arousal (LAIRD 1984). In general, the effects have been most striking for exaggeration instructions: producing an expression that is more intense than the circumstances warrant usually adds to the intensity of the experienced emotion, but the inverse effect of suppressing emotional expression is weaker and less reliable (COLLIER 1985). The reasons for this disparity are not clear, but the result is consistent with the assumption that the feedback relationship between expression and affect is mainly positive.

Studies of Aggressive Catharsis

The classical ethologists considered aggressive behavior to be a consummatory act, bearing the same relationship to aggressive motivation that eating bears to hunger (LORENZ 1950). In the absence of an opportunity for aggressive display, the level of aggressive motivation was assumed to increase continuously. Aggressive behavior was, thus, seen as cathartic because either observing or participating in an aggressive interaction could serve to release the accumulated aggressive drive (LORENZ 1966). This implies that aggressive behavior should have a negative feedback relationship to arousal: Performance should depress the likelihood of a subsequent display, rather than increasing aggressive motivation. The concept of

behavioral efference thus contrasts strongly with the classical view of aggressive motivation.

A great deal of effort has been expended in the attempt to demonstrate the cathartic effects of viewing or engaging in aggressive actions, but little experimental support has emerged. In general, vicariously experienced aggression, whether through sporting events (ARMS, RUSSELL & SANDILANDS 1979) or the media (GORANSON 1970), seems more consistent with a hypothesis of arousal than with catharsis. Observing aggressive behavior generally seems to increase, rather than reduce, aggressive feelings (BERKOWITZ 1970).

There is some support for the possibility that fully realized verbal or physical aggression can be cathartic (FESHBACH 1984), but most social scientists now attribute this effect to the release of subjects from social inhibitions against aggressive expression, rather than a release of pent-up aggressive drives (BERKOWITZ 1962, 1970; QUANTY 1976). In QUANTY's (1976) words, "When situational restraints against aggression are lowered by various experimental procedures, . . . expression of aggression leads to *increased* rather than decreased hostility on postaggression measures," a result consistent with the behavioral efference model.

Operant Conditioning Studies

Testing for the interaction of behavior with motivation in animals is more difficult than in humans because self-report measures are not available. Some relevant information can, however, be obtained from operant conditioning studies in which the opportunity to perform species-typical aggressive displays has served as the reinforcer. In a number of different species, this design has produced significant conditioning of an otherwise neutral operant (HOGAN & ROPER 1978; POTEAL 1979; HUNTINGFORD & TURNER 1987). Some of the best work has been performed with Siamese fighting fish, *Betta splendens*, in which the opportunity to engage in aggressive displays is sufficiently rewarding that it can readily be transferred to secondary reinforcers, such as a light that had previously been paired with the presentation of an opponent (THOMPSON 1969; HOGAN & ROPER 1978).

It is difficult to be certain that it is the opportunity for aggression that is reinforcing in these preparations. Some other feature of the reward situation, such as an exposure to novel stimuli, an increase in general activity, or simply an opportunity to associate with a conspecific, could be responsible for the effect (HUNTINGFORD & TURNER 1987). Careful experimental studies using a wide range of stimuli have eliminated most of these alternatives, however (BOLS 1977). The extensive literature showing appetitive behavior for aggression even in field situations also lends support to the interpretation of aggression as reinforcing (HINDE 1970).

The crucial point for the behavioral efference hypothesis is that if aggressive behavior is reinforcing for an arbitrary operant, then it must be *self*-reinforcing under circumstances of free elicitation (POTEAL 1979): Like virtue, aggression is

its own reward. BAENNINGER (1974) has drawn this same inference and generalized from it to an independent statement of the behavioral efference hypothesis:

"If the operant, or emitted, response, which is reinforced by the opportunity to behave aggressively, is itself an attack or threat response, then the probability that the aggressive response will occur again is increased simply by virtue of the fact that it is associated with performance of an aggressive response. Aggressive responses would then be self-reinforcing. Each aggressive response would be positively reinforced by its own performance in a kind of positive feedback loop" (1974, p. 25).

These studies thus constitute some of the best currently available evidence for the behavioral efference model.

Discussion

The behavioral efference hypothesis provides a feasible resolution for the game-theory paradox. Displays afford a truthful rendering of underlying aggressive motivation, at least in part because the performance of the display is instrumental in producing the realized motivational level. Revealing information about "intentions" is not selected against because the principal function of display is not communication but regulation of the rate of aggressive arousal. It may be disadvantageous to reveal one's intentions but it is far worse to forego the benefits of a graded arousal process. The need for secrecy is only one of the selective processes acting on aggressive behavior, and it may not be the most important one (COLGAN 1988; TURNER & HUNTINGFORD 1986).

Display without Communication

In addition to resolving the paradox, behavioral efference also offers a useful perspective on several other, otherwise anomalous, behavioral phenomena including displays that occur outside of normal aggressive contexts. Aggressive behavior often deviates from the idealized concept of a reciprocal, nonverbal conversation (SMITH 1977). Animals occasionally perform aggressive displays even in the absence of conspecifics. They may also display in a group of conspecifics, none of whom appears to be attending to the behavior. Even when there is a designated recipient for the display, a surprisingly large proportion of supposedly communicative actions produce no detectable response.

SMITH (1977) attributes such observations to the potentiating or "primer" functions of displays. Despite their unresponsiveness, recipients are presumed to be absorbing and integrating information about the displaying individual that will ultimately influence their long-term interactions. Whether animals do integrate such information in any meaningful fashion over the long term is difficult to assess. What is clear, however, is that if displays serve to increase aggressive arousal their performance in the absence of fully attentive recipients is understandable.

A highly aggressive individual requires only a trivial eliciting stimulus to produce a display. Even just the presence of another conspecific may suffice. The

production of a display, in its turn, feeds back to maintain the aggressive state. Even in the absence of reciprocal displays, then, behavioral efference in highly-motivated animals may be enough to maintain a continuous, erratic output of aggressive displays, in a process akin to the "run-on" or "ticking over" of a poorly tuned gasoline engine.

The Principle of Antithesis

Submissive displays commonly contain features that appear to be derived from a simple inversion of some aspect of the aggressive posture of the species (MORRIS 1956). DARWIN (1872) termed this the Principle of Antithesis, noting that such behavior patterns could only have evolved their communicatory significance by virtue of "being in complete opposition or antithesis to the attitude and movements which, from intelligible causes, are assumed when [an animal] intends to fight, and which consequently are expressive of anger" (1872, p. 51).

DARWIN's principle appears to provide an explanation for more features of submissive postures than the alternative notions that submissive displays involve exposing some vulnerable part of the body (LORENZ 1966) or concealing structures used in attack or threat (TINBERGEN & MOYNIHAN 1952; MORRIS 1956). The hypothesis has, however, generated little research, possibly because its evolutionary basis is obscure. DARWIN himself was not clear on the mechanism, attributing antithetical displays to the "unconscious" performance, through "habit and association," of actions of a directly opposite kind (DARWIN 1872).

One of the attractive features of behavioral efference is that it provides a coherent evolutionary context for the development of antithetical displays. In animals that were compelled to remain in close proximity to aggressive conspecifics there would be strong selection for mechanisms that suppressed aggressive arousal, thereby reducing the likelihood of periodic flare-ups. Submissive behavior, by opposing the actions that express aggression, opposes the positive motivational feedback that leads, through the cascading of motivation and display behavior, to a physical attack.

An animal performing actions that are fully antithetical to aggressive displays is, therefore, inhibiting the acceleration of its own aggressive predispositions. Such displays can, thus, be thought of as an additional level of control on the course of an aggressive interaction. This concept is reminiscent of CHANCE's (1962) theory of "cut-off" acts, in which an attack response is suppressed by behavior that involves looking away from the other individual.

Submissive displays will, of course, also reduce an individual's value as an aggressive stimulus to his opponent. As with the communicatory effects of aggressive displays, however, this cannot have been the principal selective influence. Without a direct connection between the physiology of aggression and its expression in the animal's behavior, submissive displays are also paradoxical because they are open to exploitation by deceitful individuals. Behaving submissively (e.g. by waving a white flag) to draw an opponent off-guard and leave him vulnerable to a surprise attack is an old and reliable tactic in human warfare.

Approaches for Further Research

Quantitative experimental investigation of behavioral efference poses several formidable problems. Aggressive displays are most readily elicited through staged encounters with an appropriate opponent. The behavior of the opponent inevitably provides a confounding source of stimulation and inhibition, however, which generally precludes a coherent test of the motivational effects of the subject's behavior.

A variety of methods involving models or puppets have been employed in ethological studies to control and manipulate the eliciting stimuli for aggression. Such simulations only roughly approximate the stimulus configuration provided by a real opponent and generally lack the flexibility necessary to reproduce a range of display intensities (but see AMLANER & STOUT 1978). The use of film images of a displaying animal could potentially yield precise stimulus control, but the technical difficulties are considerable (JENSSEN 1970).

A second problem for ethological studies of behavioral efference is the need to measure the subject animal's motivation independently of its display behavior. This has been accomplished in studies of human facial efference by measurement of autonomic responses (EKMÄN et al. 1983), and similar techniques may be feasible in animal preparations (KLING et al. 1979; WALLETSCHKE & RAAB 1982).

The most promising approach seems to be offered by neurophysiological techniques for chronic implantation of stimulatory electrodes in freely interacting animals. Aggressive behavior of varying intensities can often be reliably evoked in these subjects by varying the imposed electric current (KLING 1986). Several studies have been performed in which aggressive behavior was electronically manipulated in social contexts (DELGADO 1963, 1969; ROBINSON et al. 1969). HACCOU et al. (1988) recently explored social interactions in electrically stimulated Norway rats, using Markov analysis for separating out the effects of an opponent's behavior. An extension of their design could readily be used to quantify the effects of display production on arousal.

One feasible protocol would make use of the stimulation assay pioneered by von HOLST & VON SAINT PAUL (1963). In this design, stimulatory electrodes would be implanted in the brain of a subject animal. It would then be exposed to an appropriate eliciting stimulus, either a model or a conspecific opponent. The amount of direct electrical stimulation required to obtain attack behavior would then be measured. In the comparison treatment, the subject would first be allowed to perform aggressive displays and then subsequently would be subjected to direct electrical stimulation. If behavioral efference is present, the displaying subjects should require significantly less electrical stimulation to induce attack.

Acknowledgments

A number of people have made helpful suggestions or steered me toward useful references. I would particularly like to acknowledge the assistance of J. DIAMOND, P. COLGAN, C. FINDLAY, C. BEER, R. FRANCIS, and one anonymous reviewer.

Literature Cited

- AMLANER, C. J., & J. F. STOUT, 1978: Aggressive communication by *Larus glaucescens*, part VI. Interactions of territory residents with a remotely controlled, locomotory model. *Behaviour* 66, 223—251.
- ANDERSSON, M., 1980: Why are there so many threat displays? *J. Theor. Biol.* 86, 773—781.
- ANDREW, R. J., 1963: Evolution of facial expression. *Science* 142, 1034—1041.
- ARCHER, J., 1988: *The Behavioural Biology of Aggression*. Cambridge Univ. Press, Cambridge.
- ARMS, R. L., G. W. RUSSELL, & M. L. SANDILANDS, 1979: Effects on the hostility of spectators of viewing aggressive sports. *Soc. Psych. Quart.* 42, 275—279.
- BAENNINGER, R., 1966: Waning of aggressive motivation in *Betta splendens*. *Psychon. Sci.* 4, 241—242.
- , 1974: Some consequences of aggressive behavior: a selective review of the literature on other animals. *Aggr. Behav.* 1, 17—37.
- BARLOW, G. W., & T. E. ROWELL, 1984: The contribution of game theory to animal behavior. *Behav. Brain Sci.* 7, 101—103.
- BERKOWITZ, L., 1962: *Aggression: A Social Psychological Analysis*. McGraw-Hill, New York.
- , 1970: Experimental investigations of hostility catharsis. *J. Cons. Clin. Psych.* 35, 1—7.
- BOLS, R. J., 1977: Display reinforcement in the Siamese fighting fish, *Betta splendens*: Aggressive motivation or curiosity? *J. Comp. Physiol. Psychol.* 91, 233—244.
- BOND, A. B., 1989: Toward a resolution of the paradox of aggressive displays: I. Optimal deceit in the communication of fighting ability. *Ethology* 81, 29—46.
- BROWN, J. L., 1964: The integration of agonistic behavior in the Steller's jay, *Cyanocitta stelleri* (Gmelin). *Univ. of Calif. Pub. Zool.* 60, 223—328.
- CARYL, P. G., 1979: Communication by agonistic displays: what can games theory contribute to ethology? *Behaviour* 68, 136—169.
- , 1982a: Telling the truth about intentions. *J. Theor. Biol.* 97, 679—689.
- , 1982b: Animal signals: a reply to Hinde. *Anim. Behav.* 30, 240—244.
- CHANCE, M. R. A., 1962: An interpretation of some agonistic postures; the role of "cut-off" acts and postures. *Symp. Zool. Soc. London* 8, 71—89.
- COLGAN, P. W., 1988: *Animal Motivation*. Chapman and Hall, London.
- COLLIER, G., 1985: *Emotional Expression*. Erlbaum Assoc. Hillsdale.
- DARWIN, C., 1872: *The Expression of the Emotions in Man and Animals*. John Murray, London.
- DAWKINS, R., & J. R. KREBS, 1978: Animal signals: information or manipulation? In: *Behavioural Ecology*. (KREBS, J. R., & N. B. DAVIES, eds.) Blackwell, Oxford, pp. 282—309.
- DELGADO, J. M. R., 1963: Cerebral heterostimulation in a monkey colony. *Science* 141, 161—163.
- , 1969: Offensive-defensive behavior in free monkeys and chimpanzees induced by radio stimulation of the brain. In: *Aggressive Behaviour*. (GARATTINI, S., & E. B. SIGG, eds.) Excerpta Medica Found., Amsterdam, pp. 109—119.
- DOW, M., A. W. EWING, & I. SUTHERLAND, 1976: Studies on the behaviour of cyprinodont fish. III. The temporal patterning of aggression in *Aphyosemion striatum* (Boulenger). *Behaviour* 85, 1—24.
- EIBL-EIBESFELDT, I., 1972: Similarities and differences between cultures in expressive movements. In: *Non-Verbal Communication*. (HINDE, R. A., ed.) Cambridge Univ. Press, Cambridge, pp. 297—314.
- EKMAN, P., 1980: *The Face of Man*. Garland Press, New York.
- , 1985: *Telling Lies*. Norton, New York.
- , R. W. LEVENSON, & W. V. FRIESEN, 1983: Autonomic nervous system activity distinguishes among emotions. *Science* 221, 1208—1210.
- FESHBACH, S., 1984: The catharsis hypothesis, aggressive drive, and the reduction of aggression. *Aggr. Behav.* 10, 91—101.
- GALLUP, G. G., 1968: Mirror-image stimulation. *Psychol. Bull.* 70, 782—793.
- GEIST, V., 1974: On fighting strategies in animal combat. *Nature* 250, 354.
- GORANSON, R. E., 1970: Media violence and aggressive behavior: a review of experimental research. *Adv. Exp. Soc. Psych.* 5, 1—31.
- HACCOU, P., M. R. KRUK, E. MEELIS, E. T. VAN BAVEL, K. M. WOUTERSE, & W. MEELIS, 1988: Markov models for social interactions: analysis of electrical stimulation in the hypothalamic aggression area of rats. *Anim. Behav.* 36, 1145—1163.

- HEILIGENBERG, W., 1965: The effect of external stimuli on the attack readiness of a cichlid fish. *Z. vergl. Physiol.* **49**, 459—464.
- HINDE, R. A., 1970: *Animal Behaviour: A Synthesis of Ethology and Comparative Psychology*. McGraw-Hill, New York.
- , 1981: Animal signals: ethological and games-theory approaches are not incompatible. *Anim. Behav.* **29**, 535—542.
- , 1985: Was 'the expression of the emotions' a misleading phrase? *Anim. Behav.* **33**, 985—992.
- HOGAN, J. A., & T. J. ROPER, 1978: A comparison of the properties of different reinforcers. *Adv. Study Behav.* **8**, 155—255.
- HOLST, E. von, & U. von SAINT PAUL, 1963: On the functional organization of drives. *Anim. Behav.* **11**, 1—20.
- HUNTINGFORD, F., 1980: Analysis of the motivational processes underlying aggression in animals. In: *Analysis of Motivational Processes*. (TOATES, F. M., & T. R. HALLIDAY, eds.) Acad. Press, New York, pp. 341—356.
- , & A. TURNER, 1987: *Animal Conflict*. Chapman and Hall, London.
- IZARD, C. E., 1977: *Human Emotions*. Plenum Press, New York.
- JENSSEN, T. A., 1970: Female response to filmed displays of *Anolis nebulosus* (Sauria, Iguanidae). *Anim. Behav.* **18**, 640—647.
- KLING, A. S., 1986: The anatomy of aggression and affiliation. In: *Emotion: Theory, Research, and Experience*, Vol. 3. (PLUTCHIK, R., & H. KELLERMAN, eds.) Acad. Press, New York, pp. 237—264.
- , H. D. STEKLIS, & S. DEUTSCH, 1979: Radiotelemetered activity from the amygdala during social interactions in the monkey. *Exp. Neurol.* **66**, 88—96.
- LAIRD, J. D., 1984: The real role of facial response in the experience of emotion: a reply to Tourangeau and Ellsworth, and others. *J. Pers. Soc. Psych.* **47**, 909—917.
- LEONG, C.-Y., 1969: The quantitative effect of releasers on the attack readiness of the fish *Haplochromis burtoni* (Cichlidae, Pisces). *Z. vergl. Physiol.* **65**, 29—50.
- LEYHAUSEN, P., 1956: Verhaltensstudien bei Katzen. *Z. Tierpsychol., Beih.* **2**.
- LORENZ, K., 1950: The comparative method in studying innate behaviour patterns. *Symp. Soc. Exp. Biol.* **4**, 221—268.
- , 1966: *On Aggression*. Methuen, London.
- , 1981: *The Foundations of Ethology*. Springer-Verlag, New York.
- MAYNARD SMITH, J., 1972: Game theory and the evolution of fighting. In: *On Evolution*. Edinburgh Univ. Press, Edinburgh, pp. 8—28.
- , 1974: The theory of games and the evolution of animal conflicts. *J. Theor. Biol.* **47**, 209—221.
- , 1982: Do animals convey information about their intentions? *J. Theor. Biol.* **97**, 1—5.
- , 1984: Game theory and the evolution of behaviour. *Behav. Brain Sci.* **7**, 95—101.
- MORRIS, D. 1956: The function and causation of courtship ceremonies. *Fond. Singer-Polignac: Coll. Int. L'Instinct*, June 1954.
- , 1957: 'Typical intensity' and its relation to the problem of ritualization. *Behaviour* **11**, 1—13.
- MOYNIHAN, M., 1982: Why is lying about intentions rare during some kinds of contests? *J. Theor. Biol.* **97**, 7—12.
- OTTE, D., 1974: Effects and functions in the evolution of signalling systems. *Ann. Rev. Ecol. Syst.* **5**, 385—417.
- PLOOG, D., 1986: Biological foundations of the vocal expressions of emotions. In: *Emotion: Theory, Research, and Experience*, Vol. 3. (PLUTCHIK, R., & H. KELLERMAN, eds.) Acad. Press, New York, pp. 173—197.
- POTEGAL, M., 1979: The reinforcing value of several types of aggressive behavior: A review. *Aggr. Behav.* **5**, 353—373.
- , & L. TEN BRINK, 1984: Behavior of attack-primed and attack-satiated female golden hamsters (*Mesocricetus auratus*). *J. Comp. Psychol.* **98**, 66—75.
- QUANTY, M., 1976: Aggression catharsis. In: *Perspectives on Aggression*. (GEEN, R. G., & E. C. O'NEAL, eds.) Acad. Press, New York, pp. 99—132.
- ROBINSON, B. W., M. ALEXANDER, & G. BOWNE, 1969: Dominance reversal resulting from aggressive response evoked by brain telestimulation. *Physiol. Behav.* **4**, 749—752.

- SCHERER, K. R., 1985: Vocal affect signaling: a comparative approach. *Adv. Study Behav.* **15**, 189—244.
- SERPELL, J. A., 1982: Factors influencing fighting and threat in the parrot genus *Trichoglossus*. *Anim. Behav.* **30**, 1244—1251.
- SMITH, W. J., 1977: *The Behavior of Communicating*. Harvard Univ. Press, Cambridge.
- , 1986: An "informational" perspective on manipulation. In: *Deception: Perspectives on Human and Nonhuman Deceit*. (MITCHELL, R. W., & N. S. THOMPSON, eds.) SUNY Press, Albany, pp. 71—86.
- THOMPSON, T., 1969: Aggressive behaviour of siamese fighting fish: analysis and synthesis of conditioned and unconditioned components. In: *Aggressive Behavior*. (GARATTINI, S., & E. B. SIGG, eds.) Excerpta Med. Found., Amsterdam, pp. 15—31.
- TINBERGEN, N., 1951: *The Study of Instinct*. Clarendon Press, Oxford.
- , 1959: Comparative studies of the behaviour of gulls (Laridae): a progress report. *Behaviour* **15**, 1—70.
- , & M. MOYNIHAN, 1952: Head-flagging in the black-headed gull: its function and origin. *Brit. Birds* **45**, 19—22.
- TOATES, F., 1986: *Motivational Systems*. Cambridge Univ. Press, Cambridge.
- TOMKINS, S. S., 1962: *Affect, Imagery, Consciousness: Vol. 1. The Positive Affects*. Springer, New York.
- , 1984: Affect theory. In: *Approaches to Emotion*. (SCHERER, K. R., & P. EKMAN, eds.) Erlbaum Assoc., Hillsdale, pp. 163—195.
- TURNER, G. F., & F. A. HUNTINGFORD, 1986: A problem for game theory analysis: Assessment and intention in male mouthbrooder contests. *Anim. Behav.* **34**, 961—970.
- VAN RHIJN, J. G., 1980: Communication by agonistic displays: a discussion. *Behaviour* **74**, 284—293.
- , & R. VODEGEL, 1980: Being honest about one's intentions: an evolutionary stable strategy for animal conflicts. *J. Theor. Biol.* **85**, 623—641.
- WALLETSCHKE, H., & A. RAAB, 1982: Spontaneous activity of dorsal raphe neurons during defensive and offensive encounters in the tree-shrew. *Physiol. Behav.* **28**, 697—705.
- WEINRICH, J. D., 1980: Toward a sociobiological theory of the emotions. In: *Emotion: Theory, Research, and Experience, Vol. 1*. (PLUTCHIK, R., & H. KELLERMAN, eds.) Acad. Press, New York, pp. 113—138.
- WILZ, K. J., 1970: Causal and functional analysis of dorsal pricking and nest activity in the courtship of the three-spined stickleback *Gasterosteus aculeatus*. *Anim. Behav.* **18**, 115—124.
- ZAJONC, R. B., 1985: Emotion and facial efference: a theory reclaimed. *Science* **228**, 15—21.

Author's address: Alan B. BOND, San Diego Museum of Natural History, Balboa Park, P.O. Box 1390, San Diego, California 92112, U.S.A.